Aditya Tadimeti

adityatadimeti.com | adityatadimeti@gmail.com | linkedin.com/in/adityatadimeti/ | github.com/adityatadimeti | (650)-862-5251

Education

Stanford University — Stanford, CA
B.S. in Computer Science: Artificial Intelligence Track — *June 2025*
M.S. in Computer Science: Artificial Intelligence Track — *Sept 2025*

Work Experience

Research Engineer @ Liquid AI | PyTorch, Python — June 2025 —
- Training efficient foundation models.

Research Engineer Intern @ Adobe Firefly | PyTorch, Python — June 2024 — September 2024
- Trained text-to-image, multimodal reward models. Fine-tuned diffusion models via RLHF.

Intern of Technical Staff @ Cohere | Python, PyTorch — January 2024 — April 2024
- Worked on model fine-tuning, data ingestion, evaluation metrics, and pretraining to improve reasoning in LLMs.

SWE Intern @ Amazon | JavaScript, React, Java, AWS — June 2023 — September 2023
- Developed end-to-end full-stack service to automate supply-chain network graph cost updates.

SWE Intern @ Oracle OCI | Java, Docker, JavaScript — June 2022 — September 2022
- Built and deployed full-stack, internal debugging tool to resolve customer networking issues.

Publications

LFM2 Technical Report — November 2025
... **Aditya Tadimeti***, et al.
33 authors in alphabetical order
arXiv preprint, 2025

Simple, Scalable Reasoning via Iterated Summarization — July 2025
Vivek Vajipey*, **Aditya Tadimeti***, Justin Shen*, Ben Prystawski, Michael Y. Li, Noah Goodman
* denotes equal contribution
ICML 2025 Workshop on Long Context Foundation Models
ICML 2025 Workshop on AI for Math

Organizations

Stanford Computation & Cognition Lab | Adv Prof. Noah Goodman — April 2024 — June 2025
- Designed novel compression algorithms to optimize LLM reasoning performance at test-time (ICML '25).

Stanford SNAP Group | Adv Prof. Jure Leskovec. — September 2024 — December 2024
- Pretrained 100M+ parameter protein foundation models as an ML systems engineer on H100 cluster.

Stanford NLP Group | NLP Researcher — February 2023 — March 2024
- Researched the capability of LLMs to leverage linguistic information for prediction.

Stanford CS Department | Senior Section Leader — April 2022 — December 2024
- TA for CS 106A/B (Python/C++).

Selected Projects

Language Model from Scratch | PyTorch — Spring 2025
- Designed, implemented, and aligned Transformer language models from scratch, using only PyTorch primitives. Included custom tokenizers, model architecture, optimizers, Triton kernels, robust data processing, systems profiling, and RLHF methods like DPO and GRPO for reasoning.

AI Agent Projects | Python, React, JavaScript — 2024 - 2025
- Implemented arXiv Deep Research Discord bot that responds to research queries by traversing arXiv, multimodal AI-powered tutor that ingests text, videos, and pdfs, and AI-powered health advisor.

Quantized, Pruned, Accelerated, and Parallelized GPT-2 | PyTorch, C++, OpenMP, ISPC — Fall 2023
- Implemented 8-bit quantized inference, iterative magnitude based pruning and speculative decoding for GPT-2.
- Optimized Attention via blocking, fusing, vectorization, and multithreading; implemented FlashAttention on NanoGPT.

Coursework

ML & AI: Building Language Models from Scratch, Machine Learning, NLP, Computer Vision, ML w/ Human Preferences
Math & Statistics: Linear Algebra, Matrix Theory, Multivariable Calculus, Probability, Statistical Inference, Info. Theory
Complexity Theory: Data Structures, Algorithms, Mathematical Computing
Systems & Security: Operating Systems, Networking, Parallel Computing, Systems for Machine Learning, Web Applications

Technical Skills

Languages / Frameworks: Python, PyTorch, C++, CUDA, React, C, Java, R, JavaScript, HTML/CSS